Course Mempelajari Machine Learning dengan Python dibawakan Instruktur Bernama Raka Ardhi. Course ini berisikan 5 chapter yakni Prologue (Chapter 1), Mempersiapkan Data (Chapter 2), Memilih Algoritma (Chapter 3), Melatih Model (Chaper 4), dan Menguji Akurasi Model (Chapter 5).

**Chapter 1 – Prologue**

1. **Apa itu Machine Learning**

Machine Learning memberikan kemampuan kepada computer untuk belajar. Membangun model dari contoh input untuk membuat data-driven prediction. Machine Learning tidak membutuhkan traditional programming seperti if, case, while/for, ML membutuhkan Data, Algoritma, Analisis, dan Model. 2 Tipe ML yakni Supervised dan Unsupervised. Contoh supervised adalah prediksi harga rumah. Data => Algorithm => Model => Prediction. Contoh unsupervised adalah data cluster. Supervised : Value Prediction, Perlu hasil prediksi, model prediksi harga data baru, sedangkan Unsupervised : Data clustering, tidak perlu cluster, dan akses data cluster

1. **Persiapan**

Jupyter Notebook, Python 3, dan Anaconda 3.7.

1. **Workflow dari Machine Learning**

Permasalahan => Persiapan Data => Memilih Algoritma => Melatih Model => Uji Model

**Chapter 2 – Mempersiapkan Data**

1. **Mendapatkan Data**

Tidy Data (Rapih) = Lebih mudah dimanipulasi, variabelnya adalah kolom, observasinya adalah baris. Sumber bisa dari Google, Data Pemerintahan, dan Data Perusahaan.

1. **Load, Clean, Inspect Data**

Ada di file Kode Machine Learning.ipynb.

1. **Merapihkan Data**

Ada di file Kode Machine Learning.ipynb

**Chapter 3 – Memilih Algoritma**

1. **Algoritma Machine Learning**

Peran Algoritma. Ada beberapa factor yakni Learning Type (Supervised/Unsupervised), Result (Regression : Continue Value, prediksi tidak tetap, ex: berapa harga rumah / Classification : Diskrit Value, prediksi tetap, ex : diabetes atau tidak), Complexity, dan Basic vs Enhanced

1. **Memilih Algoritma berdasarkan kasus**

Kandidat Algoritma : Naïve Bayes (belajar dari data histori), Logistic Regression (pakai bobot), dan Decision Tree (seperti if/else, menggunakan percabangan node). Naïve Bayes : Probability Based, Bobot fitur sama, membutuhkan sedikit data, mudah dipahami, dan cukup stabil

**Chapter 4 – Melatih Model**

1. **Mengenal Model Training**

Data melatih algoritma ML, menghasilkan model, new data better prediction

1. **Proses Training**

Split data (70% Training & 30% Testing) => Train Model => Evaluasi Model. Tidak bisa testing pake data training, maka diperlukan untuk di split.

1. **Training dengan Python**

Menggunakan fungsi dari scikit-learn : data splitting, pre-processing, feature selection (mengambil minimal feature), dan model training. <= toolset scikit-learn

1. **Training Algoritma**

Ada di file Kode Machine Learning.ipynb

**Chapter 5 – Menguji Akurasi Model**

1. **Mengenal Uji Akurasi**

Menggunakan data testing, digunakan untuk meningkatkan performa sampai akurat

1. **Mengevaluasi Model Naïve Bayes**

Ada di file Kode Machine Learning.ipynb

1. **Peningkatan Performa**

Ada di file Kode Machine Learning.ipynb

1. **Apa itu Cross Validation**

Digunakan untuk validasi, bermasalah kalo datanya dikit. Bisa dipake di banyak algoritma, tinggal tambahin CV dibelakangnya, contohnya logisticregressionCV